Deep intronic sequence variants in COL2A1 affect the alternative splicing efficiency of Exon 2, and may confer a risk for rhegmatogenous retinal detachment

Spickett, Carl; Hysi, Pirro; Hammond, Chistopher J.; Prescott, Alan; Fincham, Gregory S.; Poulson, Arabella V.

Published in:
Human Mutation

DOI:
10.1002/humu.23050

Publication date:
2016

Document Version
Peer reviewed version

Human Mutation

Deep Intronic Sequence Variants in COL2A1 Affect the Alternative Splicing Efficiency of Exon 2, and May Confer a Risk for Rhegmatogenous Retinal Detachment.

Journal:	*Human Mutation*
Manuscript ID	Draft
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	n/a
Complete List of Authors:	Spickett, Carl; University of Cambridge, Pathology; University of Cambridge, Medical Genetics Hysi, Pirro; King's College London, Department of Twin Research and Genetic Epidemiology Hammond, Chris; King's College London, Department of Twin Research and Genetic Epidemiology Prescott, Alan; University of Dundee, College of Life Sciences Fincham, Gregory; Addenbrookes Hospital, Ophthalmology Poulson, Arabella; Addenbrooke's Hospital, Ophthalmology McNinch, Annie; University of Cambridge, Pathology Richards, Allan; University of Cambridge, Pathology Snead, Martin; Addenbrookes Hospital, Vitreoretinal Service
Key Words:	Alternative Splicing, COL2A1, Retinal Detachment, Posterior Vitreous Detachment

Deep Intronic Sequence Variants in *COL2A1* Affect the Alternative Splicing Efficiency of Exon 2, and May Confer a Risk for Rhegmatogenous Retinal Detachment.

Carl Spickett[1,2], Pirro Hysi[3], Chistopher J Hammond[3], Alan Prescott[4], Gregory S Fincham[1], Arabella V Poulson[1], Annie M McNinch[1,2,5], Allan J Richards[1,2,5], Martin P Snead.[1]

1. Vitreoretinal Research Group, BOX 143, Addenbrooke's Hospital, Cambridge University NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, UK.

2. Department of Pathology, University of Cambridge, Tennis Court Rd, Cambridge CB2 1QP, UK.

3. Academic Unit Ophthalmology, King's College London, St Thomas' Hospital Campus, 3rd Floor South Wing Block D Westminster Bridge Road, London SE1 7EH

4. College of Life Sciences, University of Dundee, Dow Street, Dundee, DD1 5EH, Scotland, UK

5. Regional Molecular Genetics Laboratory, BOX 143, Addenbrooke's Hospital, Cambridge University NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, UK.

Address for correspondence:

Allan Richards, Regional Molecular Genetics Laboratory, BOX 143, Addenbrooke's Hospital, Cambridge University NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, UK. e-mail:ar204@cam.ac.uk

Mr Martin P Snead, Vitreoretinal Service, BOX 41, Addenbrooke's Hospital, Cambridge University NHS Foundation Trust, Hills Road, Cambridge CB2 0QQ, UK. e-mail: mps34@cam.ac.uk

Key words: Alternative Splicing, COL2A1, Giant Retinal Tear, Posterior Vitreous Detachment, Retinal detachment

Abstract

Pathogenic *COL2A1* mutations causing haploinsufficiency of type II collagen, result in type 1 Stickler syndrome which has a high risk of retinal detachment and failure of the vitreous to develop normally. Exon 2 of *COL2A1* is alternatively spliced, expressed in the eye but not in mature cartilage and encodes a region that binds growth factors TGFβ1 and BMP-2. We investigated how both a mutation in intron 2 and a normal variant allele altered the efficiency of *COL2A1* exon 2 splicing and how the latter may act as a predisposing risk factor for the occurrence of rhegmatogenous retinal detachment (RRD) in the general population. Using both amplification of illegitimate transcripts and allele specific minigenes expressed in cultured cells, we demonstrate variability in exon 2 inclusion not only between different control individuals, but also between different *COL2A1* alleles. We also identify possible trans-acting factors that bind to allele specific RNA sequences, and investigate the effect of both knockdown and over expression of these factors on exon 2 splicing efficiency. Finally, using a specific cohort of patients with RRD and a control population, we demonstrate a significant difference in the frequency of the *COL2A1* intronic variant rs1635532 between the two groups.

Introduction

Rhegmatogenous retinal detachment (RRD; detachment due to holes or tears in the retina) is a sight threatening and potentially blinding condition if timely surgical repair is not performed. Even with successful repair, visual outcomes may be impaired by delayed presentation, macular involvement or sequelae such as proliferative vitreoretinopathy or secondary glaucoma.

In the white European population the reported incidence of retinal detachment is around 1 in 10,000 per year (Thompson *et al* 2002, Kirin *et al* 2013). Although the incidence of myopia is higher in Singapore, a study of retinal detachment there (Wong et al 1999) found similar frequencies for retinal detachment in the Chinese and Malay populations but the condition was less frequent in the Indian population (0.4 in 10,000). RRD is also much rarer in the black South African population with a rate as low as 0.46 per 100,000 individual reported (Peters 1995). This variation suggests that genetic differences may be a significant predisposing risk factor for retinal detachment independent of myopia, however very few studies have explored this further. Using a series of 181 cases of retinal detachment Go *et al* (2005) examined the frequency of detachments in siblings and offspring of affected patients and found a significantly increased incidence of detachment in siblings of patients compared to a control group. Mitry *et al* (2011) also showed higher sibling recurrence risk, suggesting genetic susceptibility for both myopic and non-myopic RRD. More recently a genome wide association study of RRD cases and controls in the white European population identified an association between a variant (rs267738) in the *CERS2* gene and RRD (Kirin *et al* 2013).

RRD is also a clinical feature of Stickler syndrome(s) (MIM #18300, 604841, 614134, 614284) a connective tissue disorder that carries a high risk of retinal detachment (Snead *et al* 2011). In addition families with dominantly inherited RRD (MIM #609508) have been found

to have mutations in *COL2A1*, one of the Stickler syndrome loci (Richards *et al* 2000a, 2005), suggesting that some cases of apparently sporadic RRD and especially giant retinal tear (GRT) may in fact be due to *de novo* or familial mutations rather than more common predisposing genetic factors. To investigate this possibility we analysed a number of patients with RRD for *COL2A1* mutations, which in turn lead us to examine the role of *COL2A1* exon 2 alternative splicing in RRD. *COL2A1* exon 2 encodes a cysteine rich domain that binds TGFβ1 and BMP-2, and acts as an antagonist to these growth factors in the eye where it is expressed (Ryan *et al* 1990, Bishop *et al* 1994, Zhu *et al* 1999). Regulation of *COL2A1* exon 2 alternative splicing is partially controlled by the nuclear protein TIA-1 and its interaction with a conserved AU rich stem loop region in intron 2, just downstream (c.292+4_+41) from the end of the exon (McAlinden *et al* 2005, 2007), however an additional splicing enhancer region(s) was predicted to exist further downstream in intron 2, somewhere between the stem loop structure and c.292+380 (McAlinden *et al* 2005).

Here we examine how a *COL2A1* intron 2 mutation and an intron 2 polymorphism affect exon 2 inclusion in the mature *COL2A1* transcript, and identify potential trans-acting splicing factors which interact with these different pre-mRNA sequences. We also examine the frequency of a polymorphic variant in a specific cohort of RRD patients compared to normal controls.

Materials and Methods

Patients and controls were selected exclusively from the white European population. Studies were performed with approval of the local ethics committee (LREC 92/019 and 02/172). Informed written consent was obtained in all cases, and adhered to the tenets of the declaration of Helsinki. Only patients with horseshoe tears and retinal detachment secondary to PVD were recruited to the study group. In addition they had to meet one of three criteria:

i) Bilateral retinal detachment secondary to posterior vitreous detachment and horseshoe retinal tear.

ii) Unilateral retinal detachment and horseshoe retinal tear(s) in the fellow eye

iii) Unilateral retinal detachment in an individual, with parents or siblings also with a horseshoe tear retinal detachment.

The control group were drawn from patients who had attended the ophthalmology department at Addenbrooke's hospital for reasons other than retinal detachment and who were over the age of 65yr such that the majority would already have experienced a PVD but without associated retinal detachment. Posterior vitreous detachment occurs in approximately 70% of the British population by the age of 65, but rarely after that age (Foos 1972, Snead *et al* 1994)

COL2A1 sequence analysis

DNA from the first 10 samples collected from the patient group was analysed as previously described for pathogenic mutations (Richards *et al* 2006). All *COL2A1* exons, and surrounding intronic regions present in the amplified region, were examined for unique sequence variants. Unique DNA variants were submitted to the LOVD *COL2A1* variant database (http://databases.lovd.nl/shared/genes/COL2A1)

RT-PCR of illegitimate transcripts

RNA obtained from cultured dermal fibroblasts was reverse transcribed with superscript II (Invitrogen), using a gene specific primer in exon 10 for the reverse transcription. The cDNA was amplified through 2 rounds of PCR using Q5 high fidelity polymerase (New England Biolabs). The primary amplification used primers in the 5'untranslated region and exon 9. The secondary PCR used primers in exon 1 and exon 8 (supplementary Table S1).

Minigene construction

Exons 1-7 of COL2A1, minus a 3164bp region of intron 1 (c.85+506-c.86-466) was amplified and cloned into the expression vector pcDNA3.1A myc-His (Invitrogen) to create minigenes with the c.292+157C>A mutation, and two wild type alleles that differed due to polymorphic variants (Table 1). The region c.1–c.85+505 was ligated to the c.86-465-c.531 region using a BamH I restriction site included in the primers used to amplify each region.

Cell culture

Cultured dermal fibroblasts, MIO-M1 Muller cell line (31) and immortalised lens epithelial cell lines (Limb *et al* 2002, Rhodes *et al* 2006) (9YRCAT, CD5A, CE11560, CE11565, CE13300 and DM1931) were maintained in DMEM (4.5 g/L glucose with 2 mM L-glutamine) with penicillin (100U/ml)/streptomycin (100µg/ml) and supplemented with non-essential amino acids, ascorbic acid (25µg/ml) and 10 % foetal bovine serum (FBS, Gibco, UK). The lens epithelial cells were obtained from either cases of congenital cataract (9YRCAT, DM1931) or from the central epithelium of donor lenses (CD5A, CE11560, CE11565, CE13300), all were transformed using the SV40 large T antigen (Rhodes *et al* 2006, Coleman *et al* 2014).

Expression of *COL2A1* minigenes and exon 2 splicing assay

Immortalised lens epithelial cells obtained from different individuals were transfected and allowed to express *COL2A1* minigenes for 24 hr before RNA was extracted using RNeasy purification kits (Qiagen UK) followed by DNase treatment using an RNase-free kit (Ambion). Minigene RNA was reverse transcribed with Superscript II (Life technologies, UK) using a vector specific primer (5'agaaggcacagtcgaggctgatcag3') and amplified using either one of two transcript-specific forward primers that bridged the junction of either exon 1 and exon 3 (5'gggccaggatgtccggcaac3') or the junction of exon 2 and 3 (5'cgccactgccagtgggcaac3') along with a reverse primer located in exon 6 (5'tttttcacctttgtcacctcgatc3'). Quantitative RT-PCR using a chromo4 real-time fluorescence detector (Biorad UK) with SYBR green PCR master mix (Biorad UK) was used to measure the relative difference between transcripts containing/lacking exon 2, using a reaction volume of 25µl and primers at a concentration of 0.24pmol/µl. After an initial denaturation of 95˚C for 3 min, amplification was achieved through 40 cycles consisting of 95˚C 30 sec, 68.5˚C 20 sec and 72˚C 15 sec. To confirm the specificity of each reaction, the amplified products were also analysed by gel electrophoresis to assess the presence of a single PCR product and its size. The amount of full length transcript (i.e. including exon 2) was calculated as a percentage of the total cDNA amplified from the same RT sample. Each amplification reaction was performed in triplicate to determine the relative difference of each transcript. This was repeated a further five times to obtain a mean difference and for statistical analysis using Student's T test.

Affinity purification (RNA direct-capture) of RNA-binding proteins and mass spectrometry

Synthetic biotinylated RNA oligonucleotides were purchased from Dharmacon (Thermo scientific; supplementary Table S2). Nuclear-enriched cell extract was purified from MIO-M1 cells using a nuclear extract kit (Active Motif). Prior to use biotinylated oligonucleotides

were conjugated to pre-equilibrated streptavidin coated M-280 magnetic beads in 1x direct capture buffer (20 mM HEPES pH 7.9, 20% glycerol, 100 mM KCl, 0.2 mM EDTA). The oligonucleotide-bead complex was incubated with pre-cleared nuclear extract, 0.5 µg of nuclear extract with 10x modified direct capture buffer without KCl (200 mM HEPES pH 7.9, 2mM EDTA, 60% glycerol) in a total volume of 500µl overnight. After washing, beads were boiled directly in modified Laemmli sample buffer containing 0.1M DTT. Proteins were separated by SDS-PAGE (10%) and visualised by silver stain using a mass-spectrometry compatible silverquest kit (Life Technologies). Liquid chromatography tandem mass spectrometry (LCI-MS/MS) and protein identification was performed on isolated and excised protein bands by the Cambridge centre for proteomics (http://proteomics.bio.cam.ac.uk/)

Western blot

Whole protein extracts from cell lines were prepared in RIPA (10 mM TRIS-HCl, 1mM EDTA, 0.5 mM EGTA, 1 % Triton X-100, 0.1 % sodium deoxycholate, 0.1 % SDS, 140 mM NaCl) buffer and protein concentration determined using a BCA (bicinchoninic acid) assay kit (Pierce). Samples were boiled in modified Laemmli sample buffer containing DTT. Proteins were separated by 10% (w/v) SDS polyacrylamide gels and transferred to nitrocellulose membranes. Membranes were blocked in 10% skimmed milk in PBST (phosphate buffered saline and 0.1% tween 20) for 1 hr at room temperature. Primary antibodies applied for western blotting analysis are catalogued in Supplementary Table S3. Western blots were visualised using LumiGLO chemiluminescent substrate (Kirkegaard & Perry Laboratories) with a GBOX chemiluminescent imager (Syngene) after probing with species-specific horseradish peroxidase conjugated secondary antibodies (Santa Cruz).

siRNA Knockdown

CE13300 cells were co-transfected with *COL2A1* reporter minigenes and small interfering RNA (siRNA; Dharmacon) using RNAi MAX (Invitrogen) followed by Lipofectamine 2000 (Invitrogen). siRNA oligonucleotides used are listed in Supplementary Table S4. Twenty four hours prior to co-transfection, cells were seeded to 1.5 x 10^6 cell density in 6 well culture dishes then transfected with 30 pmol of total siRNA using RNAi MAX. Cells were then co-transfected with 1 µg of minigene DNA and 50 pmol of siRNA with Lipofectamine 2000. RNA was extracted 48 hours post-transfection as described above. The relative abundance of the two *COL2A1* transcripts was measured using RT-PCR as described above.

Over expression of RNA binding proteins

Cells were co-transfected with vectors expressing tagged RNA binding proteins hnRNP L, DAZAP1 and TDP-43 with Lipofectamine and Lipofectamine plus reagent (Invitrogen). Briefly 2 x 10^6 cells plated in 6 well culture dishes were transfected with 0.8 µg of reporter minigene DNA and (if applicable) co-transfected with an additional 1.6 µg of plasmid expressing tagged RNA-binding proteins. RNA was isolated and *COL2A1* exon 2 expression measured as described above.

Statistical analysis

Genetic association for rs1635532 were carried out using Fisher's exact tests comparing allelic distributions in cases and controls, under the assumptions of recessive and co-dominant models for the G allele. Stata 12.1 software (College Station, TX) was used for all the calculations.

Results

COL2A1 sequence analysis

Sequencing of *COL2A1* amplified from 10 patients with RRD, identified intronic variants in three of the patients. These were c.292+157C>A in intron 2, c.3112-87delG in intron 44, and c.4318-196G>A in intron 53. Initially all three variants were absent from dbSNP (http://www.ncbi.nlm.nih.gov/snp/) and also from over 200 DNA samples sequenced as part of a *COL2A1* diagnostic sequencing service (http://www.cuh.org.uk/addenbrookes-hospital/services/genetics-laboratories). Subsequently the c.3112-87delG variant in intron 44 has been submitted to dbSNP (rs573313146), but is not validated and has no frequency data associated with it. No other unique DNA variants were detected. Mutation nomenclature is based on the human genome reference assembly hg19 and the cDNA sequence NM_001844.4

Splicing analysis

The effect of these DNA variants on splicing, of the *COL2A1* pre-mRNA, was investigated using either amplification of illegitimate transcripts from cultured fibroblasts or minigenes. No effect on splicing could be found for the c.3112-87delG and c.4318-196G>A variants and we classified these as variants with an unknown clinical significance. Illegitimate transcripts amplified using RNA from a patient's skin fibroblast cell line, with the intron 2 c.292+157C>A mutation, showed both inclusion and exclusion of the alternatively spliced exon 2 (Fig 1). Comparison to normal controls and other cases of RRD demonstrated that inclusion of exon 2 into the mature transcript was variable between samples. As the patient was also heterozygous for an intronic variant rs1635532, we genotyped the RRD and normal controls, used for amplification of illegitimate transcripts (Fig 1), to ensure a mixture of genotypes for this SNP. Interestingly all those with an rs1635532 G allele produced some

exon 2 skipping, whereas the profile of those homozygous for the A allele was variable, where two homozygous AA normal controls, produced no detectable exon skipping. It was therefore unclear if the mutation was influencing exon 2 splicing efficiency or if as suggested by the control samples other factors including the rs1635532 G allele were involved.

Exon 2 Splicing efficiency

To be able to determine if there was a quantitative difference in the efficiency of exon 2 splicing, resulting from the c.292+157C>A mutation and / or other DNA variants within the proposed splicing enhancer region, minigenes were created. These would enable differences in the efficiency of exon 2 splicing to be assessed for each *COL2A1* allele separately using the same cell line (i.e. the same genetic background), which would not be possible using endogenous illegitimate transcripts. The two *COL2A1* alleles of the patient differed in two single nucleotide polymorphisms in intron 2 (rs1793953C/T, c.292+176 and rs1635532A/G, c.292+258), that were close to the mutation, so we created a second "normal" minigene which contained the same polymorphisms as the mutant allele, except for a variable poly(A) sequence starting at c.292+264 (The haplotype of each minigene is documented in Table 1). These three minigenes were transfected into immortalised lens epithelium cell lines prepared from different individuals. RNA isolated from these transfected cells was then used with vector / transcript specific primers and real time RT-PCR to determine the ratio of *COL2A1* mRNA including and excluding exon 2, resulting from each minigene. This showed that not only was pre-mRNA from the mutant c.292+157C>A minigene less efficient at splicing exon 2 into the mature transcript but that the patient's other allele (now referred to as the C-G allele, to reflect its haplotype; Table 1) also produced less exon 2 inclusion than the other "normal" allele minigene (now referred to as the T-A allele). Minigene assays were repeated in 6 different cell lines. In all cases the C-G allele and "mutant" minigenes were less efficient

at including exon 2 in the mRNA compared to the T-A allele minigene. However splicing efficiencies varied between cell lines (Fig 2).

These results suggested that not only was exon 2 splicing variable between individuals but that other factors, in addition to the intron 2 DNA variants also contribute to exon 2 splicing efficiency. In an attempt to identify possible trans-acting splicing factors, which interact with the different alleles, we performed direct capture using biotinylated RNA oligonucleotides. In addition to the c.292+157C>A mutation, we chose to use the rs1635532 polymorphism for capture experiments, as RRD is not a common disorder and the rs1635532 G allele was less frequent (18%) in the European population than the rs1793953 C polymorphism (35%). In addition the rs1635532 G allele was more common in our cohort of RRD patients (see below), also *in silico* analysis (http://www.umd.be/HSF3/index.html) predicted that the rs1635532 alleles potentially had different affinities for trans-acting splicing factors, whereas no difference was predicted for the alleles of rs1793953. The other polymorphic SNP rs1793957 present in the minigenes was outside of the region predicted to contain additional splicing enhancer motifs (McAlinden *et al* 2005).

RNA Direct Capture

RNA direct-capture experiments were performed to identify trans-acting factors binding to the rs1635532 region within intron 2. Direct capture was performed with biotinylated RNA oligonucleotides representing the G and A alleles of rs1635532 with MIO-M1 cell nuclear extract. Mass spectrometry identified heterogenous nuclear ribonucleoprotein (hnRNP) protein family members hnRNP L and hnRNP A1 as proteins which bound to the G allele oligonucleotide and DAZ-associated protein 1 (DAZAP1) as a protein which bound with

greater affinity to the A allele oligonucleotide (Fig 3A). These were confirmed by direct capture followed by western blotting (Fig 3B).

Direct capture was repeated with biotinylated RNA oligonucleotides representing the mutant and wild type sequence of intron 2 c.292+157C>A. RNA-binding protein TDP-43 was identified by mass spectrometry, and Western blot analysis confirmed that TDP-43 had a greater affinity for the mutant (A) oligonucleotide compared to the wild-type (C) oligonucleotide (Fig 3A and 3B).

Effect of under and over expression of RNA binding proteins on COL2A1 exon 2 splicing

RNA binding proteins which showed affinity to RNA oligonucleotides by direct capture and western blotting were either under or over expressed, to investigate their effect on *COL2A1* exon 2 splicing efficiency, using minigenes expressed in CE13300 immortalised lens epithelial cells.

hnRNP A1

hnRNP A1 showed greater affinity for the rs1635532 G allele oligonucleotide (Fig 3A) and minigene assays determined that this rarer allele resulted in a decrease in exon 2 inclusion (Fig 2). siRNA depletion of hnRNP A1 rescued the exon skipping in the wild-type C-G allele minigene with exon 2 inclusion rising from 29% in luciferase siRNA controls to levels comparable to the wild-type T-A allele minigene (41%, p = <0.0001, Figure 4). A small increase was observed in depleted hnRNP A1 cells transfected with the mutant minigene compared to luciferase siRNA controls (Fig 4). hnRNP A1 over expression decreased exon 2 inclusion in both wild-type C-G allele (18 % exon 2 inclusion) and mutant minigenes (12 % exon 2 inclusion) compared to empty vector control where exon 2 inclusion for the C-G allele and mutant allele were 28 % (p = 0.0268) and 17 % (p = 0.0106) respectively (Fig 5).

TDP-43

Oligonucleotides representing the c.292+157C>A mutant and wild-type regions of *COL2A1* intron 2 identified that TDP-43 was bound with greater affinity to the mutant RNA sequence (Fig 3B). This mutation was determined to reduce exon 2 inclusion by the minigene assay (Fig 2). Compared to the luciferase control, siRNA depletion of TDP-43 rescued the exon 2 skipping observed in cells transfected with mutant minigene to a level of 42 % inclusion ($p = 0.0274$). Exon 2 inclusion was also increased significantly in TDP-43 depleted cells transfected with wild-type C-G minigene (43 %, $p = 0.0001$) but had no significant effect on splicing using the minigene of the T-A allele. Over-expression of TDP-43 (Fig 5) significantly decreased exon 2 inclusion from all three minigenes, with the wild-type T-A minigene producing only 17 % exon 2 inclusion compared to 48 % for the empty vector control ($p = 0.0002$), the wild-type C-G minigene (16 % exon 2 inclusion compared to 29 % for the empty vector control; $p = 0.0172$) and the mutant minigene (2 % exon 2 inclusion compared to 17 % for the empty vector control; $p = <0.0001$).

hnRNP L

Another hnRNP family member hnRNP L bound with greater affinity to the rs1635532 G allele oligonucleotide compared to the A allele oligonucleotide (Fig 3A). siRNA depletion of hnRNP L increased exon 2 inclusion in both the wild-type C-G allele and mutant minigenes to levels comparable to the wild-type T-A allele (Fig 4). But did not increase the level of exon 2 inclusion in RNA from the T-A allele minigene above that seen in the luciferase control. Over expression of hnRNP L significantly increased exon 2 inclusion in RNA from the wild-type T-A minigene (74 % compared to 48 % for the empty vector control; $p = 0.0005$), C-G allele minigene (40 % exon 2 inclusion compared to 29 % for the empty vector

control; p = 0.0034) and the mutant minigene (36 % exon 2 inclusion compared to 17 % for the empty vector control; p = 0.0006).

DAZAP1

Although DAZAP1 binds to both rs1635532 G and A allele oligonucleotides, it bound with greater affinity to the A allele oligonucleotide (Fig 3A). DAZAP1 depletion by siRNA (Fig 4) resulted in almost complete inclusion of exon 2 in the wild type T-A allele minigene (93 %), wild type C-G allele minigene (88 %) and the mutant minigene (89 %). Surprisingly DAZAP1 over-expression (Fig 5) also significantly increased exon 2 inclusion in the wild type T-A minigene (65 % exon 2 inclusion compared to 48 % for the empty vector control; p = 0.0172), the wild type C-G minigene (40 % exon 2 inclusion compared to 29 % for the empty vector control; p = 0.0087) and the mutant minigene (25 % exon 2 inclusion compared to 17 % for the empty vector control; p = 0.0156). However over expression did not reach the high level of exon 2 inclusion observed when expression of DAZAP1 was reduced.

Population genotyping and statistical analysis

As the rs1635532 G allele was shown to result in less efficient *COL2A1* exon 2 inclusion in the minigene system, we used a Fisher's exact test to compare distributions of the rs1635532 alleles between 244 RRD patients and 215 healthy controls, (Table 2) We observed an association between the rs1635532 alleles and RRD status. This association was strongest and most significant under a recessive model, where individuals homozygous for the G allele were more likely to be RRD cases, than subjects with at least one copy of allele A (OR=2.23, 95% CI 1.16-4.44, p=0.01), but less significant under the assumptions of a co-dominant model (OR= 1.31 95%CI 0.98-1.75, p=0.058).

Discussion

Although many "risk alleles" have been identified via genome wide association studies, it is probable that to date, only a minority of those have identified the functional DNA sequence variant that confers the risk to the disorder under investigation. Here by initially studying the effect of a mutation in *COL2A1* we also identified a functional effect of a relatively common variant rs1635532. We then showed statistically significant association between homozygosity for the rs1635532 G allele and a carefully selected cohort of patients with a specific type of retinal detachment. As both the mutation and common variant are deep within an intron it is unlikely that either of these would have been identified using exome sequencing. Indeed the rs1635532 polymorphism does not appear in the exome sequence variation database (http://evs.gs.washington.edu/EVS/), this may have implications for other studies still searching for a functional difference between alleles of genes identified in association studies.

Functional role of *COL2A1* exon 2

Alternative splicing of *COL2A1* is likely to have key regulatory roles in the development of both eye and cartilage. The cysteine rich domain encoded by exon 2 is expressed in the eye but not in mature cartilage (Sandell *et al* 1991, Bishop *et al* 1994), it can bind TGFβ1 and BMP-2 (Zhu *et al* 1999) and therefore act as an antagonist to these growth factors. The role of the type II collagen cysteine rich domain in regulating growth factor homeostasis suggests that perturbation of this function could play a role in the development of RRD. Type II collagen is the major structural component of the vitreous, and under-expression of the molecule, due to nonsense mediated decay, is the most common molecular cause of type 1 Stickler syndrome, where affected patients have a high risk of RRD, as well as high congenital myopia due to megalophthalmos, which is unusual as most neonates are born

hypermetropic. TGFβ1 homeostasis is known to be disrupted due to mutations in the extracellular matrix protein fibrillin (Neptune *et al* 2003) which result in Marfan syndrome (MIM #154700). Like Stickler syndrome, Marfan syndrome has an increased incidence of retinal detachment (Nemet *et al* 2006), suggesting that a possible common molecular mechanism (i.e. altered TGFβ1 homeostasis) accounts for the increased risk of RRD seen in these two disorders. In mature cartilage *COL2A1* exon 2 is not expressed, and in this tissue both TGFβ1and BMP-2 have been shown to activate cartilage specific gene expression (Murphy *et al* 2015) explaining why the antagonistic effect of the, exon 2 encoded, cysteine rich domain is not required in that particular tissue. This region of the type II collagen N-propeptide may also be partially retained in heterotypic type II/XI collagen heterotrimers, as it has been detected on the outside of collagen fibrils in the vitreous, where it is available for interaction with other extracellular matrix molecules or cell receptors (Reardon *et al* 2000).

Control of exon 2 alternative splicing

Regulation of *COL2A1* exon 2 splicing is in part controlled by interaction of the RNA binding protein TIA-1 with a stem loop structure present between the +4 and +41 region of intron 2 pre-mRNA sequence (McAlinden *et al* 2007). Although it has been shown that disruption of this stem loop structure diminished exon 2 inclusion, deletion of bases +11 to +380 of intron 2 completely abolished exon 2 inclusion (McAlinden *et al* 2005). This led the authors to suggest the existence of other exon 2 splicing enhancer region(s) further downstream from the stem loop region.

The mutant and polymorphic intron 2 regions, identified here as modifiers of *COL2A1* exon 2 splicing, do not necessarily pinpoint these sequences as part of the proposed additional splicing enhancers. Instead the ability of the mutant and polymorphic alleles to bind various nuclear proteins may inhibit the efficiency with which other trans-acting splicing factors can

bind to intronic splicing enhancer motifs. This effect was best demonstrated by the knockdown and over expression of hnRNP A1, where reduction in the expression of this protein lead to increased exon 2 inclusion in RNA from the minigene with the rs1635532 G allele, to which it bound specifically. Conversely over expression of hnRNP A1 reduced exon 2 inclusion from the C-G allele minigene, with little or no effect on splicing efficiency from the other two minigenes. A similar effect was seen when TDP-43 expression was altered. This protein bound preferentially to the mutant allele (but also to the normal alleles) and reduction in its expression lead to an increase in exon 2 inclusion in RNA from the mutant minigene. As it also had the same effect on RNA from the C-G allele minigene, but not the T-A allele minigene, this suggests that there is some interaction between the binding of TDP-43 at the c.292+157C/A nucleotide and hnRNP A1 that inhibits exon 2 inclusion. This also corresponds with the observation that hnRNP A1 binds almost exclusively to the rs1635532 G allele. Over expression of TDP-43 decreased exon 2 inclusion in RNA from all three minigenes but had a greater effect on the mutant RNA to which it binds preferentially.

The result of under and over expression of the RNA binding proteins hnRNP L and DAZAP1 is more difficult to explain. Particularly for DAZAP1 which increased exon 2 inclusion when both under and over expressed, compared to the control samples. Similar results were achieved with hnRNP L except under expression had little or no effect using the T-A allele minigene, whereas over expression had a much greater effect on RNA from the T-A allele minigene than the other two (C-G allele and mutant) minigenes. One explanation for these results is that altered splicing efficiency may not be as a direct consequence of binding to *COL2A1* pre-mRNA sequences, but may instead be secondary to changes in expression of other trans-acting splicing factors, caused by altered DAZAP1 and hnRNP L expression. Since it is known that expression of splicing factors can themselves be regulated by

alternative splicing and other trans-acting splicing factors (Barberan-Soler and Zahler 2008, Dreumont *et al* 2010).

Results from the minigene experiments can also account for the RT-PCR profiles obtained from amplification of illegitimate *COL2A1* transcripts using dermal fibroblast RNA. Here only very low levels of transcripts were available to be amplified. So in those individuals with an rs1635532 G allele, where efficiency of exon 2 inclusion was lower, the likelihood of amplifying exon 2 skipped cDNAs is greater than for homozygous rs1635532 A individuals.

Consequences of misregulation of alternative splicing.

Alternative splicing is a common feature of many human genes (Pan *et al* 2008) it allows tissue specific expression of different protein isoforms thus increasing the proteomic diversity expressed from the human genome. Pathogenic mutations that are present within or that affect normal splicing of alternatively spliced exons will therefore usually only affect the specific tissues in which they are expressed, or in which they have a particular function that cannot be compensated for by other isoforms. For instance, missplicing of exon 8 of the *VCAN (CSPG2)* gene, which results in Wagner syndrome (Miyamoto *et al* 2005), only affects the eye, even though exon 8 is also expressed in other tissues. Mutations in alternatively spliced exons can also result in modified phenotypes which are considerably milder than related disorders caused by mutations in constitutive exons of the same gene. Examples of this type of effect are mutations in *USH1C* that result in non syndromic hearing loss rather than Usher syndrome type 1c (Ouyang *et al* 2002), mutations in exon 9 of *COL11A1* contribute to a recessive form of type 2 Stickler syndrome instead of fibrochondrogenesis (Richards *et al* 2013), and a mutation in *BBS8* that results in retinitis pigmentosa instead of Bardet-Biedl syndrome (Riazuddin *et al* 2010). Polymorphic variants, rather than mutations,

have also recently been shown to modify the splicing efficiency of various exons in the *OPMR1*, *PCLO* and *HGMCR* genes (Seo *et al* 2013, Xu *et al* 2014, Yu *et al* 2014).

Mutations in *COL2A1* exon 2 that result in nonsense mediated decay and / or some exon skipping cause a predominantly ocular form of type 1 Stickler syndrome, with little if any systemic features of the disorder (Richards *et al* 2000b, McAlinden *et al* 2008). Therefore it might be expected that variants / mutations that merely affect the efficiency of exon 2 inclusion will result in very minor, if any, clinical phenotype. To determine if the functional difference detected here for the rs1635532 / c.292+258 G allele had an influence on a predisposition to develop retinal detachment, we genotyped a cohort of patients with RRD and a control group.

Analysis of the cohort of RRD patients (Table 2) showed that there were a significantly greater number of individuals homozygous for the rs1635532 c.292+258 G allele than there were in a control group. Interestingly the patient with the c.292+157C>A DNA change had a mutated rs1635532 A allele in addition to a G allele and so did not contribute to the group of RRD patients that were homozygous for c.292+258G. As the cohort of RRD patients have not had the complete *COL2A1* gene sequenced it is unclear what contribution new mutations, that subtly alter the expression of the gene, may make towards a predisposition towards RRD. We have certainly detected two other sequence changes in the gene (c.3112-87delG and c.4318-196G>A) that at present are either very rare or unique to the patients, but have an unknown clinical significance. If there is a high incidence of new mutations that contribute to RRD, then this will affect the ability to detect association between any particular common gene variant and a predisposition to develop RRD, as the mutations could occur *de novo* on any allele.

Although a previous genome wide association study had not indicated that variants of *COL2A1* were a predisposing risk factor for RRD (Kirin *et al* 2013), there may be reasons for

this. Firstly in the case of rare disorders such as retinal detachment, association studies are severely disadvantaged by the difficulty of recruitment of sufficient subject numbers. Attempts to bolster numbers of affected individuals risks degrading any observable association by amalgamating different phenotypic sub-groups of retinal detachment, which potentially have a different underlying molecular mechanism. For this reason we limited our cohort of patients to those resulting from horseshoe retinal tears secondary to pathological posterior vitreous detachment and did not include RRD due to round holes, retinal dialysis or retinoschisis. Secondly as we have demonstrated here, new mutations in addition to common DNA variants may predispose individuals to RRD.

Despite the relatively low numbers of individuals tested, by focussing on a defined specific sub-group of RRD, we found a statistically significant association between the rs1635532 c.292+258 G allele and RRD associated with posterior vitreous detachment (PVD). Clearly this analysis needs to be replicated with other populations in addition to increasing the numbers of patients to confirm this association. However the link between the functional consequence of the polymorphism, subsequent effect on TGFβ1 and BMP-2 homeostasis, and a predisposition towards RRD is persuasive, and is reinforced by the identification of the intron 2 mutation which had a similar effect to the polymorphism.

Other factors must also contribute to the risk of this type of RRD, as many of the normal control group were also homozygous for the rs1635532 c.292+258 G allele. It is possible that these other factors may modify the effect of the rs1635532 G allele. If they also vary between populations, than this may explain why the rs1635532 G allele appears to be more common in the African population than Europeans, even though the incidence of RRD in Africans is apparently lower (the frequency of the rs1635532 alleles in the African population is only based on 46 chromosomes; http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=1635532). The current work has

also identified potential candidates for these factors, namely molecules involved in controlling splicing of exon 2, such as TIA-1 hnRNP A1, and TDP-43, which can potentially modify the effect of the rs1635532 G allele. Other candidates include TGFβ1, BMP-2 and factors that may, like the *COL2A1* exon 2 encoded region, regulate their action. Now that it is possible to cost effectively sequence the complete *COL2A1* gene (~30kb), by next generation sequencing, it should be possible to determine the contribution that both polymorphic variation in these candidate genes and mutations in *COL2A1* make to the development of RRD. If individuals at a high risk of RRD could be identified, in particular those with a family history of giant retinal tear, then prophylactic cryotherapy to reduce the risk of detachment, as described for treating patients with Stickler syndrome may be a possibility (Fincham *et al* 2014).

Acknowledgements

We would like to thank Professor Chris Smith for use of the hnRNP L construct and for helpful discussion with members of his lab, Dr Franco Pagani for DAZAP1 and hnRNP A1 constructs and Professor Jerne Ule for the TDP-43 construct.

We thank the Cambridge centre for proteomics for their expertise and mass spectrometry analysis of RNA direct capture samples.

We thank Mrs Wieslawa Johnson for help with the patient recruitment.

This work was supported by the University of Cambridge Retinal Research Fund and the Annie Arnold Research legacy

References

Barberan-Soler S, Zahler AM. 2008. Alternative splicing regulation during C. elegans development: splicing factors as regulated targets. PLoS Genet 4:e1000001.

Bishop PN, Reardon AJ, McLeod D, Avad S. 1994. Identification of alternatively spliced variants of type II procollagen in vitreous. Biochem Biophys Res Commun 203:289-295.

Coleman SM, Prescott AR, Sleeman JE. 2014. Transcriptionally correlated subcellular dynamics of MBNL1 during lens development and their implication for the molecular pathology of myotonic dystrophy type 1. Biochem J 458:267-280.

Dreumont N, Hardy S, Behm-Ansmant I, Kister L, Branlant C, Stévenin J, Bourgeois CF. 2010. Antagonistic factors control the unproductive splicing of SC35 terminal intron. Nucleic Acids Res 38:1353-1366.

Fincham GS, Pasea L, Carroll C, McNinch AM, Poulson AV, Richards AJ, Scott JD, Snead MP. 2014. Prevention of retinal detachment in Stickler syndrome: the Cambridge prophylactic cryotherapy protocol. Ophthalmology 121:1588-1597.

Foos RY 1972. Posterior vitreous detachment. Trans Am Acad Ophthal and Otol 76:480-497.

Go SL, Hoyng CB, Klaver CCW. 2005. Genetic risk of rhegmatogenous retinal detachment. Arch Ophthalmol 123:1237-1241.

Kirin M, Chandra A, Charteris DG, Hayward C, Campbell S, Celap I, Bencic G, Vatavuk Z, Kirac I, Richards AJ, Tenesa A, Snead MP, *et al* 2013. Genome-wide association study identifies genetic risk underlying primary rhegmatogenous retinal detachment. Hum Mol Genet 22:3174-3185.

Limb GA, Salt TE, Munro PM, Munro PM, Moss SE, Khaw PT. 2002. In vitro characterization of a spontaneously immortalized human Muller cell line (MIO-M1). Invest Ophthalmol Vis Sci 43:864-869.

McAlinden A, Havlioglu N, Liang L, Davies SR, Sandell LJ. 2005. Alternative splicing of type II procollagen exon 2 is regulated by the combination of a weak 5' splice site and an adjacent intronic stem-loop cis element. J Biol Chem 280:32700-32711.

McAlinden A, Liang L, Mukudai Y, Imamura T, Sandell LJ. 2007. Nuclear protein TIA-1 regulates COL2A1 alternative splicing and interacts with precursor mRNA and genomic DNA. J Biol Chem 282:24444-24454.

McAlinden A, Majava M, Bishop PN, Perveen R, Black GC, Pierpont ME, Ala-Kokko L, Männikkö M. 2008. Missense and nonsense mutations in the alternatively-spliced exon 2 of COL2A1 cause the ocular variant of Stickler syndrome. *Hum Mutat* 29:83-90

Mitry D, Williams L, Charteris DG, Fleck BW, Wright AF, Campbell H. 2011. Population-based estimate of the sibling recurrence risk ratio for rhegmatogenous retinal detachment. Invest Ophthalmol Vis Sci 52:2551-2555.

Miyamoto T, Inoue H, Sakamoto Y, Kudo E, Naito T, Mikawa T, Mikawa Y, Isashiki Y, Osabe D, Shinohara S, Shiota H, Itakura M. 2005. Identification of a novel splice site mutation of the CSPG2 gene in a Japanese family with Wagner syndrome. Invest Ophthalmol Vis Sci 46:2726-2735.

Murphy MK, Huey DJ, Hu JC, Athanasiou KA. 2015. TGF-β1, GDF-5, and BMP-2 stimulation induces chondrogenesis in expanded human articular chondrocytes and marrow-derived stromal cells. Stem Cells 33:762-773.

Nemet AY, Assia EI, Apple DJ, Barequet IS. 2006. Current concepts of ocular manifestations in Marfan syndrome. Surv Ophthalmol 51:561-575.

Neptune ER, Frischmeyer PA, Arking DE, Myers L, Bunton TE, Gayraud B, Ramirez F, Sakai LY, Dietz HC. 2003. Dysregulation of TGF-beta activation contributes to pathogenesis in Marfan syndrome. Nat Genet 33:407-411.

Ouyang X-M, Xia X-J, Verpy E, Du L-L, Pandya A, Petit C, Balkany T, Nance WE, Liu, X-Z. 2002. Mutations in the alternatively spliced exons of USH1C cause non-syndromic recessive deafness. Human Genetics 111:26-30.

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nat Genet 40:358-363.

Peters AL. 1995. Retinal detachment in black South Africans. S Afr Med J 85:158-159.

Reardon A, Sandell L, Jones CJ, McLeod D, Bishop PN. 2000. Localization of pN-type IIA procollagen on adult bovine vitreous collagen fibrils. Matrix Biol 19:169-173.

Rhodes JD, Monckton DG, McAbney JP, Prescott AR, Duncan G. 2006. Increased SK3 expression in DM1 lens cells leads to impaired growth through a greater calcium-induced fragility. Hum Mol Genet 15:3559-3568.

Riazuddin SA, Iqbal M, Wang Y, Masuda T, Chen Y, Bowne S, Sullivan LS, Waseem, NH, Bhattacharya S, Daiger SP, Zhang K, Khan SN, *et al.* 2010. A splice-site mutation in a retina-specific exon of BBS8 causes nonsyndromic retinitis pigmentosa. Am J Hum Genet 86:805-812.

Richards AJ, Baguley DM, Yates JR, Lane C, Nicol M, Harper PS, Scott JD, Snead MP. 2000a. Variation in the vitreous phenotype of Stickler syndrome can be caused by different amino acid substitutions in the X position of the type II collagen Gly-X-Y triple helix. Am J Hum Genet 67:1083-1094.

Richards AJ, Martin S, Yates JRW, Scott JD, Baguley DM, Pope FM Snead MP. 2000b. COL2A1 exon 2 mutations: relevance to the Stickler and Wagner syndromes. Br J Ophthalmol 84:364-371.

Richards AJ, Meredith S, Poulson A, Bearcroft P, Crossland G, Baguley DM, Scott JD, Snead MP 2005. A novel mutation of COL2A1 resulting in dominantly inherited rhegmatogenous retinal detachment. Invest Ophthalmol Vis Sci 46:663-668.

Richards AJ, Laidlaw M, Whittaker J, Treacy B, Rai H, Bearcroft P, Baguley DM, Poulson A, Ang A, Scott JD, Snead MP. 2006. High efficiency of mutation detection in type 1 stickler syndrome using a two-stage approach: vitreoretinal assessment coupled with exon sequencing for screening COL2A1. Hum Mutat 27:696-704.

Richards AJ, Fincham GS, McNinch A, Hill D, Poulson AV, Castle B, Lees MM, Moore, AT, Scott JD, Snead MP. 2013. Alternative splicing modifies the effect of mutations in COL11A1 and results in recessive type 2 Stickler syndrome with profound hearing loss. J Med Genet 50:765-771.

Ryan MC, Sandell LJ. 1990. Differential expression of a cysteine-rich domain in the amino-terminal propeptide of type II (cartilage) procollagen by alternative splicing of mRNA. J Biol Chem 265:10334-10339.

Sandell L, Morris N, Robbins JR, Goldring MB. 1991. Alternatively spliced type II procollagen mRNAs define distinct populations of cells during vertebral development: differential expression of the amino-propeptide. J Cell Biol 114:1307-1319.

Seo S, Takayama K, Uno K, Ohi K, Hashimoto R, Nishizawa D, Ikeda K, Ozaki N, Nabeshima T, Miyamoto Y, Nitta A. 2013. Functional analysis of deep intronic SNP rs13438494 in intron 24 of PCLO gene. PLoS One 8:e76960.

Snead MP, Snead DRJ, Mahmood AS, Scott JD. 1994. Vitreous detachment and the posterior hyaloid membrane: A clinicopathological study. Eye 8:204-209.

Snead MP, McNinch AM, Poulson AV, Bearcroft P, Silverman B, Gomersall P, Parfect V, Richards AJ. 2011. Stickler syndrome, ocular only variants and a key diagnostic role for the ophthalmologist. Eye 25:1389-1400.

Thompson JA, Snead MP, Billington BM, Barrie T, Thompson JR, Sparrow JM. 2002. National audit of the outcomes of primary surgery for rhegmatogenous retinal detachment (I) Samples and methods. Eye 16:766-770.

Wong TY, Tielsch JM, Schein OD. 1999. Racial difference in the incidence of retinal detachment in Singapore. Arch Ophthalmol 117:379-383.

Xu J, Lu Z, Xu M, Pan L, Deng Y, Xie X, Liu H, Ding S, Hurd YL, Pasternak GW, Klein RJ, Cartegni L, *et al*. 2014. A heroin addiction severity-associated intronic single nucleotide polymorphism modulates alternative pre-mRNA splicing of the μ opioid receptor gene OPRM1 via hnRNPH interactions. J Neurosci 34:11048-11066.

Yu CY, Theusch E, Lo K, Mangravite LM, Naidoo D, Kutilova M, Medina MW. 2014. HNRNPA1 regulates HMGCR alternative splicing and modulates cellular cholesterol metabolism. Hum Mol Genet 23:319-332.

Zhu Y, Oganesian A, Keene DR, Sandell LJ. 1999. Type IIA procollagen containing the cysteine-rich amino propeptide is deposited in the extracellular matrix of prechondrogenic tissue and binds to TGF-β1 and BMP-2. J Cell Biol 144:1069-1080.

Figure Legends

Figure 1. Amplification of illegitimate COL2A1 transcripts from cultured dermal fibrobalsts.

RNA from cultured skin fibroblasts obtained from cases of RRD (RD) including a patient (P) with a *COL2A1* intron 2 mutation and normal controls (N) was reverse transcribed and subjected to two rounds of amplification by PCR, using nested primers for the second amplification reaction. Amplified DNA was visualised by agarose gel electrophoresis along with size standards (S) where the 500bp standard is indicated. The individuals were also genotyped for the SNP rs1635532 to ensure that at both homozygous and heterozygous individuals for this variant were included as controls.

Figure 2. Effect of *COL2A1* DNA variants on exon 2 splicing in the minigene system. Expression of *COL2A1* minigenes in lens epithelial cell lines followed by qRT-PCR measured exon 2 inclusion. The graph is a quantification of six replicates expressed as means ± SEM. In 9YRCAT cells exon 2 inclusion with the T-A allele minigene was 41% compared to 11% with the C-G allele minigene and 20% with the mutant T-A allele minigene. In CD5A cells the corresponding levels of exon 2 inclusion was 34%, 20% and 15% respectively. In CE11560 cells 40%, 21% and 22% respectively. In CE11565 cells 31%, 8% and 19 % respectively. In CE13300 cells 46%, 25% and 16% respectively. In DM1931 cells 57%, 28% and 30% respectively. Levels of exon 2 inclusion which significantly differed between the different minigenes for each cell line tested are indicated with an asterisk (* = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$).

Figure 3. Direct-capture of RNA-binding proteins.

Silver stained PAGE of proteins captured by RNA oligonucleotides for the G and A alleles of rs1635532 (A) or the mutant c.292+157 C>A (B). The position of size standards (kD) loaded onto on the gel (but not shown) are indicated. Mass spectrometry identified DAZAP1 as a captured protein that appeared to have a greater affinity to the rs1635532 A-allele oligonucleotide. Similarly hnRNP A1 and hnRNP L were identified as proteins with affinity only to the G allele oligonucleotide. This was confirmed by western blot analysis, where hnRNP D, which was not identified by mass spectrometry, served as a negative control for the western blot along with a bead only (no oligonucleotide) control for the direct capture. TDP-43 was captured by both mutant and wild-type oligonucleotides, however with greater affinity to the mutant oligonucleotide, as determined by western blotting. Although SF1 was identified by mass spectrometry as a protein that appeared to be captured by the mutant oligonucleotide, western blotting did not confirm this difference. Although identified by mass spectrometry hnRNP K was not identified by western blot, and so no further analysis was undertaken with these two proteins. hnRNP M and hnRNP Q served as negative controls along with a bead only control for the direct capture.

Figure 4. The effect of siRNA depletion of RNA-binding proteins on *COL2A1* exon 2 splicing.

The cell line CE11330 was pre-treated with siRNAs for RNA binding proteins as indicated and then transfected with one of three minigenes for different alleles of *COL2A1* exons 1-7 (T/A =T-A allele, C/G= C-G allele, Mt = mutant T-A allele). The amount of exon 2 inclusion in mRNA transcribed from each minigene was measured by real time RT-PCR and compared to RNA from cells pre-treated with a luciferase siRNA control. Levels of exon 2 inclusion which significantly differed from the luciferase control are indicated with an asterisk (* = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$; see also results section). Reduction in expression of

each RNA binding protein was confirmed by western blotting lysates from the luciferase siRNA transfected cells with lysates from cells co-transfected with siRNA targeting a specific RNA-binding protein and minigene as indicated. Equal loading was determined by blotting for beta-actin.

Figure 5. Effect of RNA-binding protein over-expression on *COL2A1* exon 2 splicing.

Over-expression of hnRNP A1 significantly decreases exon 2 inclusion in mRNA from the wild-type C-G allele and mutant minigenes. Both hnRNP L and DAZAP1 over-expression significantly increase exon 2 inclusion in mRNA from all three *COL2A1* minigenes. Over-expression of TDP-43 significantly decreases exon 2 inclusion in mRNA from all three *COL2A1* minigenes. The results are a quantification of six replicates expressed as means ± SEM. Levels of exon 2 inclusion which significantly differed from the empty vector control are indicated with an asterisk (* = $p<0.05$, ** = $p<0.01$, *** = $p<0.001$). Over-expression of each RNA binding protein was confirmed by western blotting using lysates from the non-transfected cells compared with lysates from cells co-transfected with an RNA-binding protein construct along with a *COL2A1* minigene. Equal loading was determined by blotting for beta-actin.

Table 1 Polymorphic differences between minigenes

Polymorphism	Normal "T-A" allele	Mutant c.292+157A	Normal "C-G" allele
rs1793953 / c.292+176	T	T	C
rs1635532 / c.292+258	A	A	G
c.292+264 Poly A region	17As	18As	20As
rs1793957 / c.293-610	C	C	G

Table 2 Distribution of rs1635532 genotypes in the RRD and control population cohorts.

NVC = Normal Vitreous Control, RRD = Rhegmatogenous Retinal Detachment

	NVCs		RRDs	
	A	G	A	G
Totals	298	132	309	175
Homozygous	98	15	102	35
Heterozygous	102	102	105	105

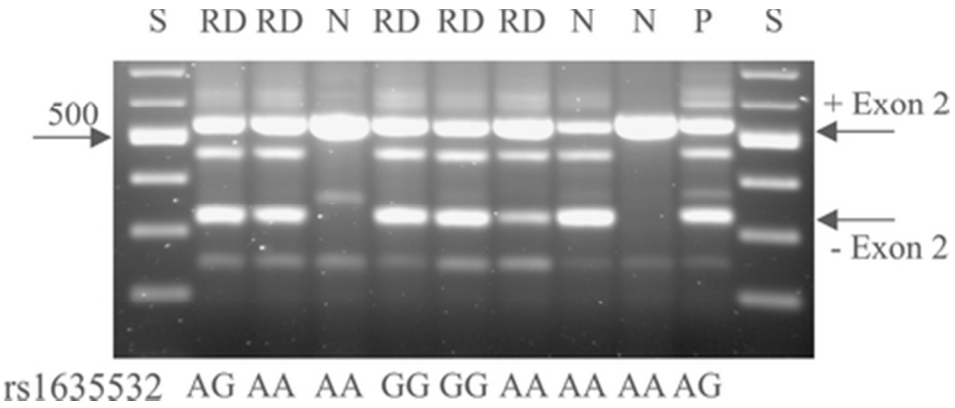

Fig 1
39x16mm (300 x 300 DPI)

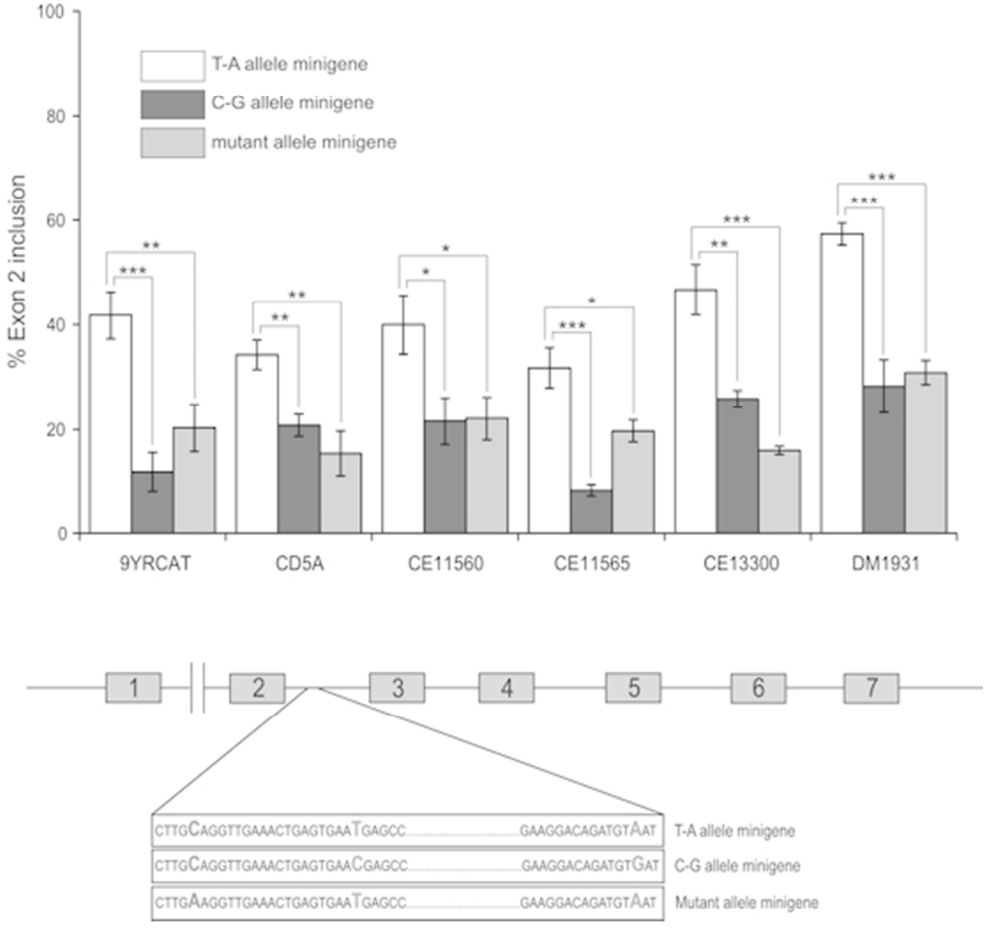

Fig 2
200x185mm (72 x 72 DPI)

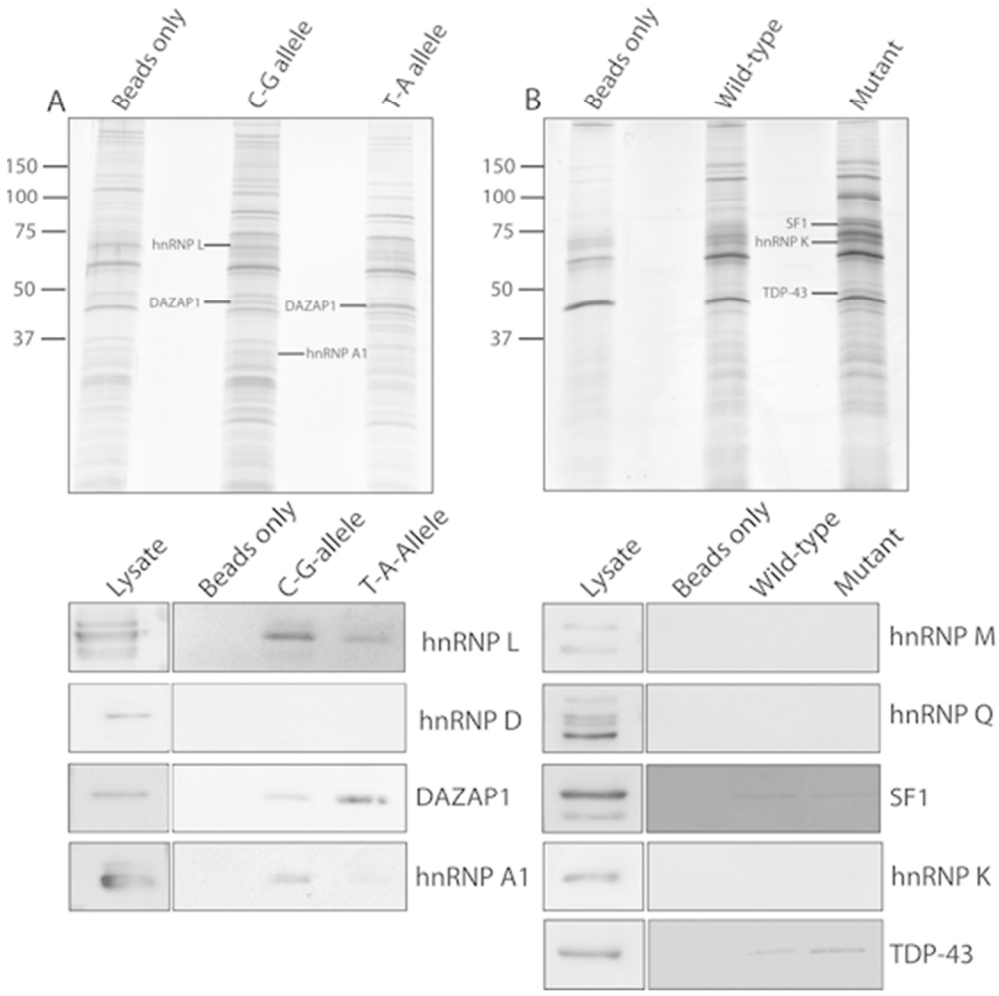

Fig 3
200x197mm (72 x 72 DPI)

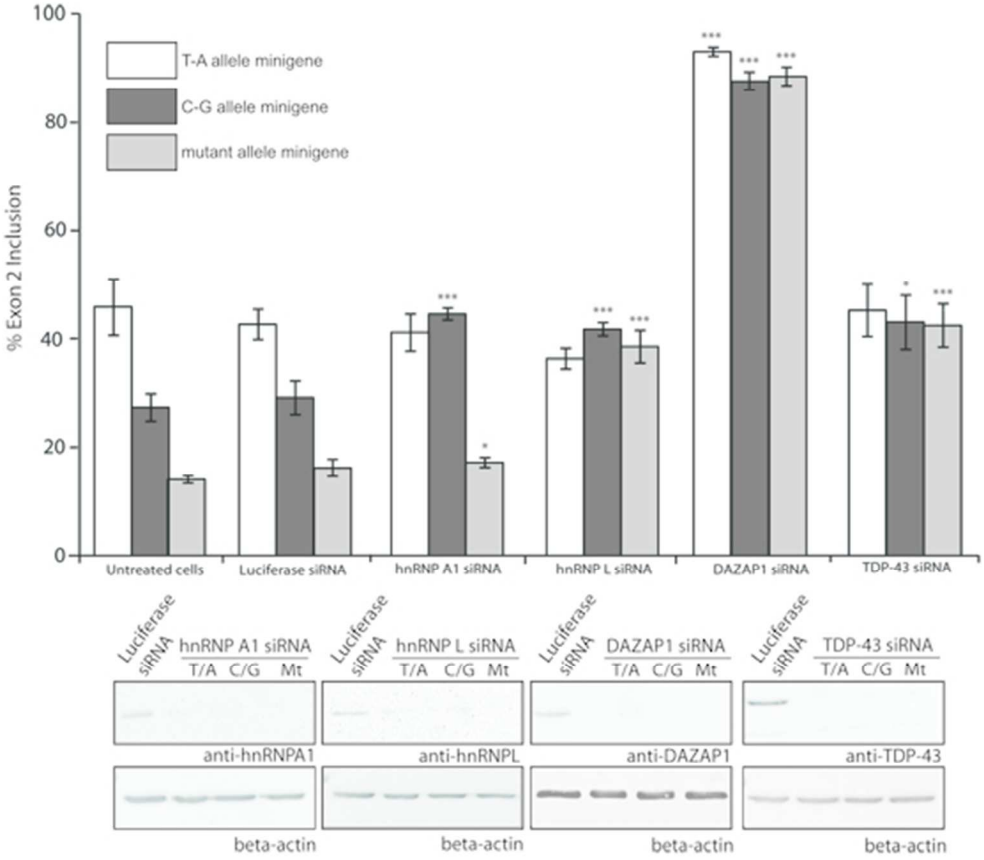

Fig 4
200x172mm (72 x 72 DPI)

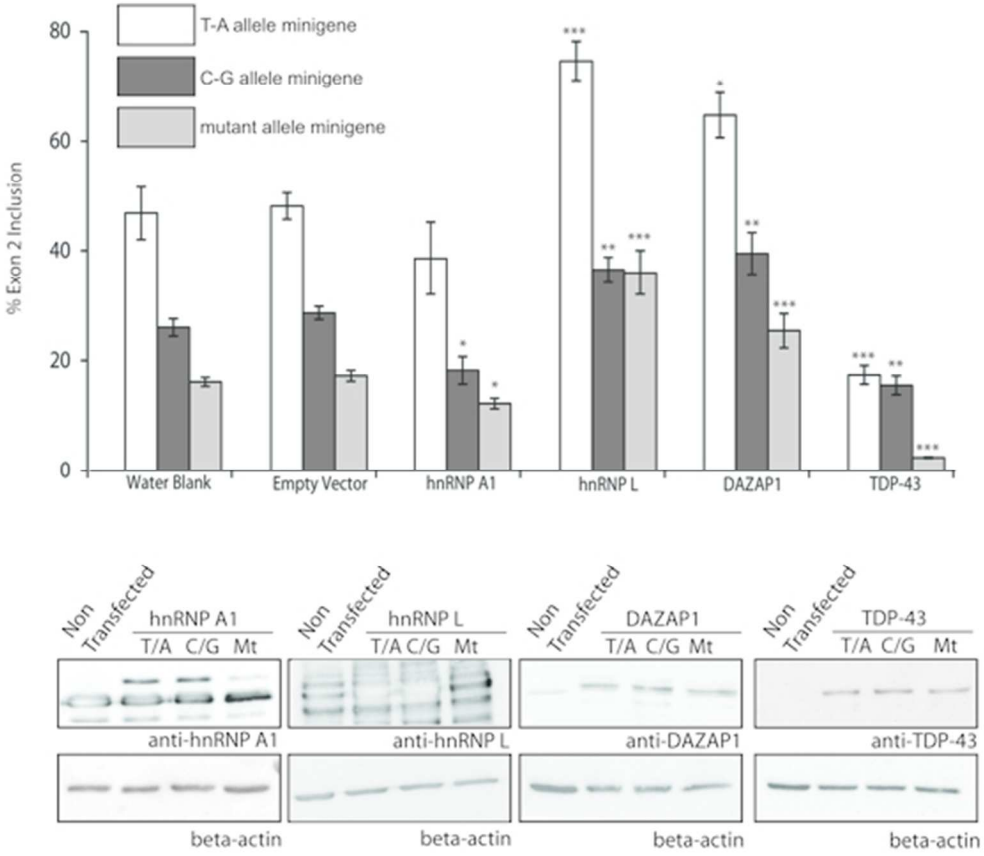

Fig 5
200x172mm (72 x 72 DPI)